

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/669,781

Source: \_\_\_\_\_

Date Processed by STIC: \_\_\_\_\_

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/669,781

DATE: 10/18/2004

TIME: 10:08:11

Input Set : N:\Crif3\RULE60\10669781.raw.txt

Output Set: N:\CRF4\10182004\J669781.raw

```

1 <110> APPLICANT: APAJALAHTI, JUHA
2   HEIKKINEN, PEKKA
3   KEROVUO, JANNE
4   LAURAEUS, MARKO
5   MORGAN, ANDREW
6   NURMINEN, PAIVI
7   SIIKANEN, OSMO
8 <120> TITLE OF INVENTION: PHYTASE FROM BACILLUS SUBTILIS, GENE
9   ENCODING SAID PHYTASE, METHOD FOR ITS PRODUCTION AND USE
10 <130> FILE REFERENCE: 65462
11 <140> CURRENT APPLICATION NUMBER: US/10/669,781
12 <141> CURRENT FILING DATE: 2003-09-24
13 <150> PRIOR APPLICATION NUMBER: US/09/242,499
14 <151> PRIOR FILING DATE: 2000-01-10
15 <150> PRIOR APPLICATION NUMBER: GB 9616957.8
16 <151> PRIOR FILING DATE: 1996-08-13
17 <150> PRIOR APPLICATION NUMBER: PCT/EP97/04385
18 <151> PRIOR FILING DATE: 1997-08-12
19 <160> NUMBER OF SEQ ID NOS: 2
20 <170> SOFTWARE: FastSEQ for Windows Version 3.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 1290
24 <212> TYPE: DNA
25 <213> ORGANISM: Bacillus subtilis
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (91)...(1239)
29 <400> SEQUENCE: 1
30   cacatttgac aattttcaca aaaacttaac actgacaatc atgtatatat gttacaattg      60
31   aagtgcacgt tcataaaagg aggaagtaaa atg aat cat tca aaa aca ctt ttg      114
32                               Met Asn His Ser Lys Thr Leu Leu
33                               1           5
34   tta acc gcg gcg gcc gga ctg atg ctc aca tgc ggt gcg gtg tct tcc      162
35   Leu Thr Ala Ala Ala Gly Leu Met Leu Thr Cys Gly Ala Val Ser Ser
36       10           15           20
37   cag gca aag cat aag ctg tcc gat cct tat cat ttt acc gtg aat gca      210
38   Gln Ala Lys His Lys Leu Ser Asp Pro Tyr His Phe Thr Val Asn Ala
39       25           30           35           40
40   gcg gcg gaa acg gaa ccg gtt gat acg gcc ggt gac gcg gct gat gat      258
41   Ala Ala Glu Thr Glu Pro Val Asp Thr Ala Gly Asp Ala Ala Asp Asp
42       45           50           55
43   cct gcg att tgg ctg gac ccc aag act cct cag aac agc aaa ttg att      306
44   Pro Ala Ile Trp Leu Asp Pro Lys Thr Pro Gln Asn Ser Lys Leu Ile

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/669,781

TIME: 10:08:11

Input Set : N:\Crf3\RULE60\10669781.raw.txt

Output Set: N:\CRF4\10182004\J669781.raw

45		60		65		70		
46	acg	acc	aat	aaa	aaa	tca	ggt	tta gtc gtt tac agc ctt gat ggt aag 354
47	Thr	Thr	Asn	Lys	Lys	Ser	Gly	Leu Val Val Tyr Ser Leu Asp Gly Lys
48		75		80		85		
49	atg	ctt	cat	tcc	tat	aat	acc	ggg aag ctg aac aat gtc gat atc cgt 402
50	Met	Leu	His	Ser	Tyr	Asn	Thr	Gly Lys Leu Asn Asn Val Asp Ile Arg
51		90		95		100		
52	tat	gat	ttt	ccg	ttg	aac	ggc	aaa aaa gtc gat atc gcg gca gca tcc 450
53	Tyr	Asp	Phe	Pro	Leu	Asn	Gly	Lys Lys Val Asp Ile Ala Ala Ala Ser
54	105			110		115		120
55	aat	cgg	tct	gaa	gga	aaa	aat	acc att gag att tac gct att gat gga 498
56	Asn	Arg	Ser	Glu	Gly	Lys	Asn	Thr Ile Glu Ile Tyr Ala Ile Asp Gly
57				125		130		135
58	aaa	aac	ggc	aca	tta	caa	agc	atg aca gat cca gac cat ccg att gca 546
59	Lys	Asn	Gly	Thr	Leu	Gln	Ser	Met Thr Asp Pro Asp His Pro Ile Ala
60		140		145		150		
61	aca	gca	att	aat	gag	gta	tac	ggt ttt acc tta tac cac agt caa aaa 594
62	Thr	Ala	Ile	Asn	Glu	Val	Tyr	Gly Phe Thr Leu Tyr His Ser Gln Lys
63		155		160		165		
64	aca	gga	aaa	tat	tac	gcg	atg	gtg aca gga aaa gag ggt gaa ttt gaa 642
65	Thr	Gly	Lys	Tyr	Tyr	Ala	Met	Val Thr Gly Lys Glu Gly Glu Phe Glu
66		170		175		180		
67	caa	tac	gaa	tta	aag	gcg	gac	aaa aat gga tac ata tcc ggc aaa aag 690
68	Gln	Tyr	Glu	Leu	Lys	Ala	Asp	Lys Asn Gly Tyr Ile Ser Gly Lys Lys
69	185			190		195		200
70	gta	cgg	gcg	ttt	aaa	atg	aat	tcc cag acg gaa ggg atg gca gca gac 738
71	Val	Arg	Ala	Phe	Lys	Met	Asn	Ser Gln Thr Glu Gly Met Ala Ala Asp
72				205		210		215
73	gat	gaa	tac	ggc	agg	ctt	tat	atc gca gaa gaa gat gag gcc att tgg 786
74	Asp	Glu	Tyr	Gly	Arg	Leu	Tyr	Ile Ala Glu Glu Asp Glu Ala Ile Trp
75		220		225		230		
76	aag	ttc	agc	gcc	gag	ccg	gac	ggc ggc agt aac gga acg gtt atc gac 834
77	Lys	Phe	Ser	Ala	Glu	Pro	Asp	Gly Gly Ser Asn Gly Thr Val Ile Asp
78		235		240		245		
79	cgt	gcc	gac	ggc	agg	cat	tta	act cgt gat att gaa gga ttg acg att 882
80	Arg	Ala	Asp	Gly	Arg	His	Leu	Thr Arg Asp Ile Glu Gly Leu Thr Ile
81		250		255		260		
82	tac	tac	gct	gct	gac	ggg	aaa	ggc tat ctg atg gca tca agc cag gga 930
83	Tyr	Tyr	Ala	Ala	Asp	Gly	Lys	Gly Tyr Leu Met Ala Ser Ser Gln Gly
84	265			270		275		280
85	aac	agc	agc	tac	gcc	att	tat	gac aga caa gga aag aac aaa tat gtt 978
86	Asn	Ser	Ser	Tyr	Ala	Ile	Tyr	Asp Arg Gln Gly Lys Asn Lys Tyr Val
87				285		290		295
88	gcg	gat	ttt	cgc	ata	aca	gac	ggt cct gaa aca gac ggg aca agc gat 1026
89	Ala	Asp	Phe	Arg	Ile	Thr	Asp	Gly Pro Glu Thr Asp Gly Thr Ser Asp
90		300		305		310		
91	aca	gac	gga	att	gac	gtt	ctg	ggt ttc gga ctg ggg cct gaa tat ccg 1074
92	Thr	Asp	Gly	Ile	Asp	Val	Leu	Gly Phe Gly Leu Gly Pro Glu Tyr Pro
93		315		320		325		

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/669,781

TIME: 10:08:11

Input Set : N:\Crif3\RULE60\10669781.raw.txt

Output Set: N:\CRF4\10182004\J669781.raw

```

94      ttc ggt att ttt gtc gca cag gac ggt gaa aat ata gat cac ggc caa      1122
95      Phe Gly Ile Phe Val Ala Gln Asp Gly Glu Asn Ile Asp His Gly Gln
96      330                      335                      340
97      aag gcc aat caa aat ttt aaa atc gtg cca tgg gaa aga att gct gat      1170
98      Lys Ala Asn Gln Asn Phe Lys Ile Val Pro Trp Glu Arg Ile Ala Asp
99      345                      350                      355                      360
100     caa atc ggt ttc cgc ccg ctg gca aat gaa cag gtt gac ccg aga aaa      1218
101     Gln Ile Gly Phe Arg Pro Leu Ala Asn Glu Gln Val Asp Pro Arg Lys
102     365                      370                      375
103     ctg acc gac aga agc gga aaa taaacatgca aaaagcagct tatacaagct      1269
104     Leu Thr Asp Arg Ser Gly Lys
105     380
106     gctttttgca tgtgaagaac g      1290
108 <210> SEQ ID NO: 2
109 <211> LENGTH: 383
110 <212> TYPE: PRT
111 <213> ORGANISM: Bacillus subtilis
112 <400> SEQUENCE: 2
113     Met Asn His Ser Lys Thr Leu Leu Leu Thr Ala Ala Ala Gly Leu Met
114     1                      5                      10                      15
115     Leu Thr Cys Gly Ala Val Ser Ser Gln Ala Lys His Lys Leu Ser Asp
116     20                      25                      30
117     Pro Tyr His Phe Thr Val Asn Ala Ala Ala Glu Thr Glu Pro Val Asp
118     35                      40                      45
119     Thr Ala Gly Asp Ala Ala Asp Asp Pro Ala Ile Trp Leu Asp Pro Lys
120     50                      55                      60
121     Thr Pro Gln Asn Ser Lys Leu Ile Thr Thr Asn Lys Lys Ser Gly Leu
122     65                      70                      75                      80
123     Val Val Tyr Ser Leu Asp Gly Lys Met Leu His Ser Tyr Asn Thr Gly
124     85                      90                      95
125     Lys Leu Asn Asn Val Asp Ile Arg Tyr Asp Phe Pro Leu Asn Gly Lys
126     100                     105                     110
127     Lys Val Asp Ile Ala Ala Ala Ser Asn Arg Ser Glu Gly Lys Asn Thr
128     115                     120                     125
129     Ile Glu Ile Tyr Ala Ile Asp Gly Lys Asn Gly Thr Leu Gln Ser Met
130     130                     135                     140
131     Thr Asp Pro Asp His Pro Ile Ala Thr Ala Ile Asn Glu Val Tyr Gly
132     145                     150                     155                     160
133     Phe Thr Leu Tyr His Ser Gln Lys Thr Gly Lys Tyr Tyr Ala Met Val
134     165                     170                     175
135     Thr Gly Lys Glu Gly Glu Phe Glu Gln Tyr Glu Leu Lys Ala Asp Lys
136     180                     185                     190
137     Asn Gly Tyr Ile Ser Gly Lys Lys Val Arg Ala Phe Lys Met Asn Ser
138     195                     200                     205
139     Gln Thr Glu Gly Met Ala Ala Asp Asp Glu Tyr Gly Arg Leu Tyr Ile
140     210                     215                     220
141     Ala Glu Glu Asp Glu Ala Ile Trp Lys Phe Ser Ala Glu Pro Asp Gly
142     225                     230                     235                     240
143     Gly Ser Asn Gly Thr Val Ile Asp Arg Ala Asp Gly Arg His Leu Thr

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/669,781

TIME: 10:08:11

Input Set : N:\Crf3\RULE60\10669781.raw.txt

Output Set: N:\CRF4\10182004\J669781.raw

144				245				250				255	
145	Arg	Asp	Ile	Glu	Gly	Leu	Thr	Ile	Tyr	Tyr	Ala	Ala	Asp
146				260				265				270	
147	Tyr	Leu	Met	Ala	Ser	Ser	Gln	Gly	Asn	Ser	Ser	Tyr	Ala
148			275				280				285		
149	Arg	Gln	Gly	Lys	Asn	Lys	Tyr	Val	Ala	Asp	Phe	Arg	Ile
150		290				295			300				
151	Pro	Glu	Thr	Asp	Gly	Thr	Ser	Asp	Thr	Asp	Gly	Ile	Asp
152		305				310			315			320	
153	Phe	Gly	Leu	Gly	Pro	Glu	Tyr	Pro	Phe	Gly	Ile	Phe	Val
154				325				330			335		
155	Gly	Glu	Asn	Ile	Asp	His	Gly	Gln	Lys	Ala	Asn	Gln	Asn
156			340				345			350			
157	Val	Pro	Trp	Glu	Arg	Ile	Ala	Asp	Gln	Ile	Gly	Phe	Arg
158		355				360			365				
159	Asn	Glu	Gln	Val	Asp	Pro	Arg	Lys	Leu	Thr	Asp	Arg	Ser
160		370				375			380				

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/669,781

DATE: 10/18/2004

TIME: 10:08:12

Input Set : N:\Crf3\RULE60\10669781.raw.txt

Output Set: N:\CRF4\10182004\J669781.raw